

A Map of the Brazilian Stock Market

Leonidas Sandoval Junior

Insper, Instituto de Ensino e Pesquisa

July 22, 2011

Abstract

We use the correlation matrix of stocks returns in order to create maps of the São Paulo Stock Exchange (BMF-Bovespa), Brazil's main stock exchange. The data refer to the year 2010, and the correlations between stock returns lead to the construction of a minimum spanning tree and of asset graphs with a variety of threshold values. The results are analysed using techniques of network theory.

1 Introduction

The BM&F-Bovespa (Bolsa de Valores, Mercadorias e Futuros de São Paulo) is the major stock exchange in Brazil, with a market capitalization of US\$ 927 million in 2010. It is also a good representative of a stock exchange of an emerging market, and has received increasing attention by international investors in the past years.

The aim of this article is to present a map, actually two maps, of the BM&F-Bovespa in 2010. In order to do so, I shall use some techniques taken from Random Matrix Theory [1], first developed for the use in nuclear physics and then used in many areas, including finance (see [2] for a comprehensive list of contributions). There are many studies of networks built from data of financial markets around the world, mainly based on the New York Stock Exchange [3]-[18], but also using data from Nasdaq [18], the London Stock Exchange [19] [20], the Tokyo Stock Exchange [19] [21], the Hong Kong Stock Exchange [19], the National Stock Exchange of India [22], the Global financial market [11] [19] [23]-[26], the USA Commodity market [27], the foreign currency market [28]-[31], and the world trade market [32]-[35]. Until the present date, to the author's knowledge, no work has been done in using the stocks of BM&F-Bovespa as a source of data for developing networks.

Both maps are done using the time series of the 190 stocks of BM&F-Bovespa that were negotiated every day the stock exchange was open (a list is given in Appendix A), so that those stocks are all very liquid. From the time series of daily prices, I obtained the series of log-returns, given by

$$S_t = \ln(P_t) - \ln(P_{t-1}) \approx \frac{P_t - P_{t-1}}{P_t} , \quad (1)$$

where P_t is the price of a stock at day t and P_{t-1} is the price of the same stock at day $t - 1$. The correlation matrix between all log-returns was then calculated using the data obtained for the whole year of 2010.

The time series of stocks prices encode an enormous amount of information about the way they relate to each other, and only part of that information is captured by the correlation matrix of their log-returns. This information also presents a good amount of noise, and should be filtered whenever that is possible. In this work, I employ some threshold values based on simulations of randomized data based on the time series that is under study in order to eliminate some of that noise. The shuffled data is obtained by reordering in a random way every time series of every single log-return. This generates time series that have the exact probability distribution as the original data, but with the connections between each log-return made completely random. The correlation matrix obtained from the randomized data is then compared with the correlation matrix of the original data. All connection values that are of comparable to the ones of randomized data are then eliminated or marked as possibly random.

There are many measures of correlation between elements of time series, the most popular being the Pearson correlation coefficient. The drawback of this correlation measure is that it only detects linear relationships between two variables. Two different measures that can detect nonlinear relations are the Spearman and the Kendall tau rank correlations, which measure the extent to which the variation of one variable affects other variable, without that relation being necessarily linear. In this work, I chose Spearman's rank correlation, for it is fairly fast to calculate and it is better at measuring nonlinear relations.

The correlation matrix may then be used to create a distance matrix, where distance is a measure of how uncorrelated two log-return series are from one another. There are also many ways to build a distance measure from the correlations between data. In this work, I shall use a different metric from [3], which is a nonlinear mapping of the Pearson correlation coefficients between stock returns. The metric to be considered here differs from the aforementioned metric because it is a linear realization of the Spearman rank correlation coefficient between the indices that are being studied:

$$d_{ij} = 1 - c_{ij} , \quad (2)$$

where c_{ij} are elements of the correlation matrix calculated using Spearman's rank correlation. This distance goes from the minimum value 0 (correlation 1) to the maximum value 2 (correlation -1).

Using the distance matrix, I shall build two maps (networks) based on the stocks of BM&F-Bovespa. The first map shall be based on a Minimum Spanning Tree (MST), which is a graph where each stock is a node (vertex), connected to one or more nodes. An MST is a planar graph with no intersections in which all nodes are connected and the sum of distances is minimum. The number of connections is the same as the total number of nodes of the network, minus one. This type of representation is very useful for visualizing connections between stocks, but it often oversimplifies the original data. Another representation is called an *asset graph*, which may be built by establishing a threshold under which correlations are considered, eliminating all other correlations above the said threshold. This also makes the number of original connections drop, simplifying the information given by the original correlation matrix. A three-dimensional map of stocks may be built in such a way that the distances portrayed in it are the best approximation to the real distances, and the connections obtained by establishing a threshold are then drawn on such representation.

The maps are examples of two networks that can be obtained from the same original data, each with its advantages and drawbacks. Using the two maps, one may then be able to identify which stocks are more dependent on each, and also which ones are more connected to others, what may be useful when building portfolios that minimize risk through diversification [36]. There are measures of how central each node of a network is, establishing its overall importance in the web of nodes, and also ways to visualize the overall distribution of those measures. As we shall see, the maps help verify that stocks that belong to the same types of companies tend to agglomerate in the same clusters, and that stocks with weak correlations often tend to connect at random in those networks.

The MST is built in section 2, its centrality measures are shown in section 3, and their cumulative distribution functions are studied in section 4. The assets graphs are built in section 5, and the centrality of one of them are studied in section 6. Section 7 discusses the k-shell decomposition of one of the asset graphs, and section 8 presents a conclusion and general discussion of results.

2 Minimum spanning trees

Minimum spanning trees are networks of nodes that are all connected by at least one edge so that the sum of the edges is minimum, and which present no loops. This kind of tree is particularly useful for representing complex networks, filtering the information about the correlations between all nodes and presenting it in a planar graph.

As discussed in the introduction, I shall employ simulations of randomized data in order to establish a distance threshold above which correlations are seen as possibly of random nature. The result of 1000 simulations of randomized data is a lower threshold 0.69 ± 0.02 , so distances above this value are represented as dashed lines in the minimum spanning tree diagrams.

The network formed by the stocks of the BM&F-Bovespa which were negotiated every day the stock exchange functioned during 2010 have six main hubs, which are displayed in figure 1. The stocks are VALE3 and VALE5,

belonging to Vale, the major mining industry in Brazil and one of the largest in the world, BBDC4, stock from Bradesco, one of the major banks in Brazil, BRAP4, which is a branch of Bank Bradesco responsible for its participations in other companies, GFSA3, stocks from GAFISA, and PDGR3, stocks from PDG Realty, both construction and materials companies.

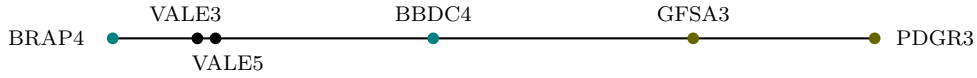


Figure 1: main axis of the minimum spanning tree for 2010.

For reasons of clarity, I shall divide the minimum spanning tree into five diagrams, or clusters, each one constructed around one of the five main hubs. The first cluster we shall study in more detail, figure 2, is the one formed around BRAP4, Bradespar, which is a company created when Bradesco (banking) was dismembered. It manages the participations of Bradesco in other, non-financial companies, particularly CPFL and Vale. As can be seen from the excerpt of the minimum spanning tree, it has strong ties with VALE3 (of Vale, a mining industry of which it has about 17% of the stocks). It is immediately surrounded by 13 stocks, three of them of financial background (PINE4 and ABCB4 are both stocks of banks, PSSA3, of an insurance company), health (TEMP3 and AMIL3), sanitation (SBSP3, itself linked with stocks of another sanitation company, CSMG3), mining (VALE3 and MMXM3), petrochemistry (GPCP3), logistics (ALLL3, LLXL2, and OHLB3), and agribusiness (SLCE3). Indirectly and weakly connected (as shown by the dashed line) to this hub are the stocks of another insurance company (SULA11). This is a somewhat mixed cluster, with a variety of stocks orbiting the stocks of a very capitalized bank (Bradesco is the third major bank of Brazil).

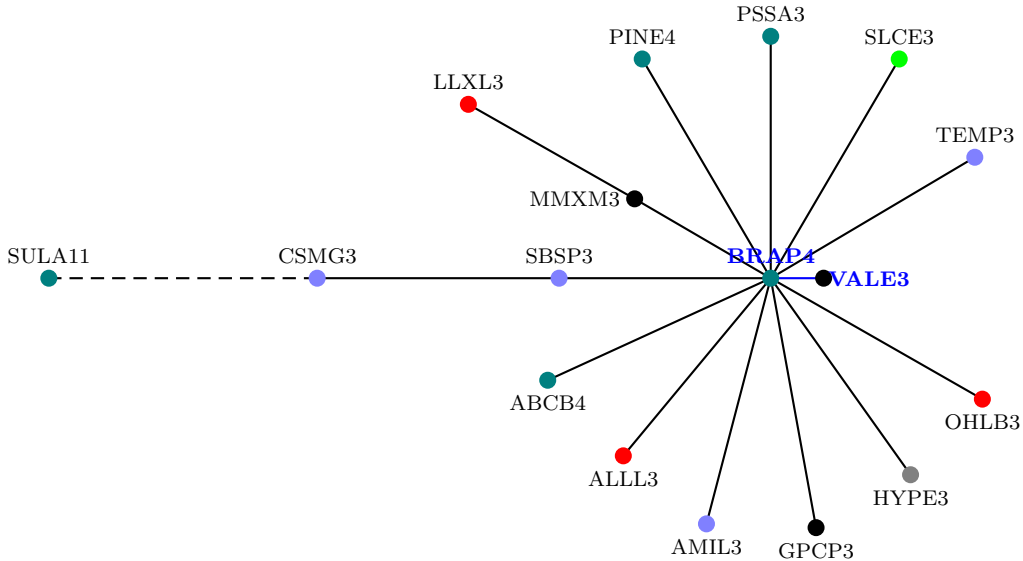


Figure 2: cluster around BRAP4.

The second and third clusters are actually the same, and they are represented in different figures in order to make it easier to discern the organization of the stocks around VALE3 and VALE5. Beginning with figure 3, one can see a cluster formed around VALE3, which is strongly correlated with VALE5, as it would be expected. Vale is a mining company, and one can observe that connected to it, on its top, there is a collection of other stocks of mining companies (CSNA3, USIM3, and USIM5), of metalurgy (GGBR3, GGBR4, and GOAU4, all of them from the Gerdau group, and MAGG3), and of petrochemistry (UNIP6). Also connected, indirectly, to VALE3, are the stocks PETR3 and PETR4, of Petrobras, an oil, gas, and biofuel company which is one of the major in the world. Petrobras is responsible for a large amount of the volume traded in the Bovespa, and it is surprising to find it not as a hub itself in the minimum spanning tree representation.

A likely explanation is that the main product of Petrobras, petrol, is a commodity negotiated worldwide and so much more dependent on factors like the price of the oil barrel than on internal ones. This is enough to place Petrobras' stocks apart from the others.

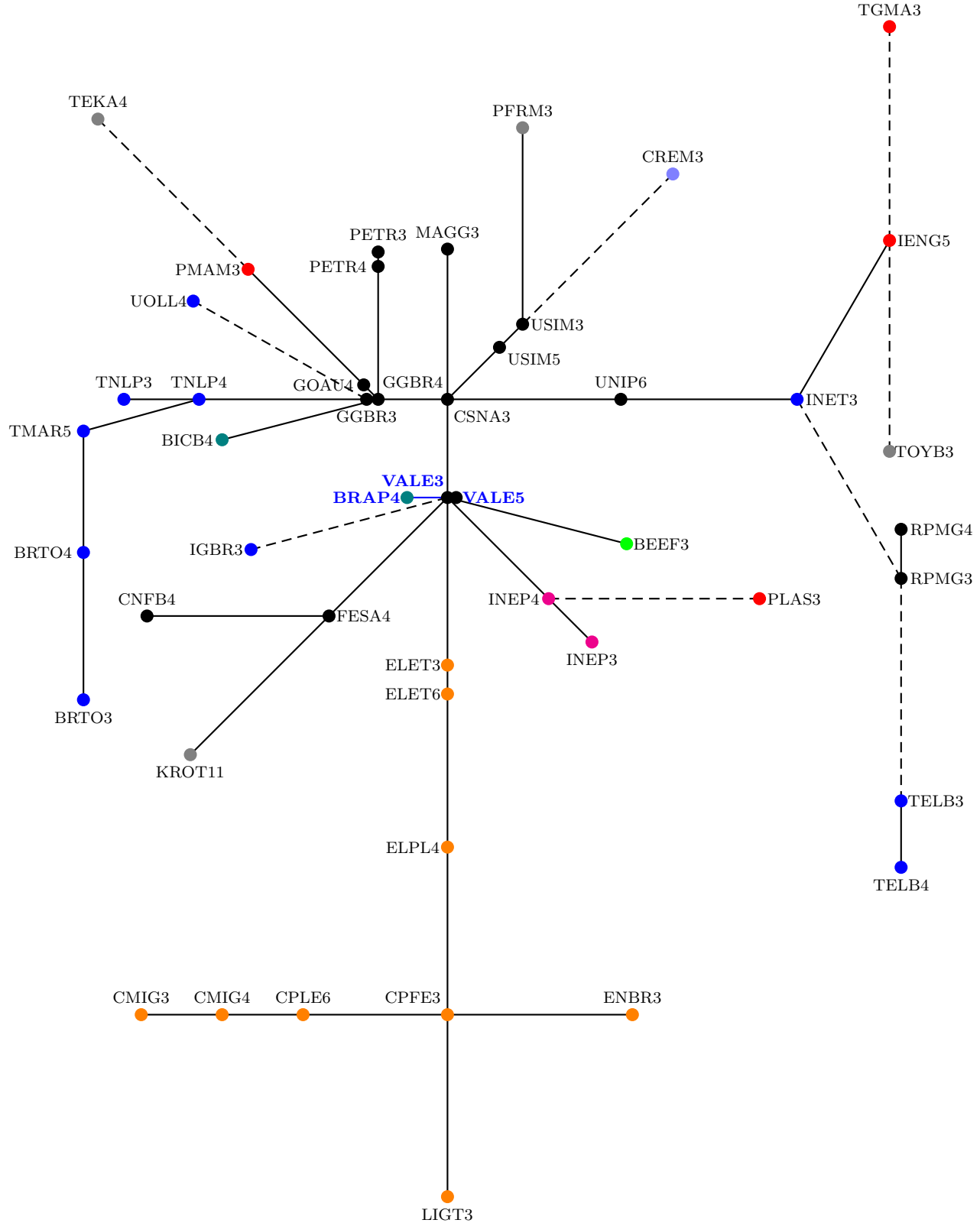


Figure 3: cluster around VALE3.

At the lower part of figure 3, one can see a cluster of stocks related with electricity distribution (ELET3 and ELET6, ELPL4, CPFE3, ENBR3, LIGT3, CPLE6, and CMIG3 and CMIG4), most sparsely correlated with

one another. To the left, there is a cluster of stocks related with telecommunications (TNPL3 and TNPL4, TMAR5, and BRTO3 and BRTO4). There are also some more stocks, related with engineering and materials, logistics, food, heavy machinery, consumer goods, and health, scattered and not forming any particular cluster. Note that most of these stocks, arranged apparently at random, have weak connections, as shown by the dashed lines.

Figure 4 shows the cluster around VALE5, which is the same cluster as the one of figure 3, since VALE3 and VALE5 are intimately connected. At the right of VALE5, one can see a cluster of stocks belonging to companies related to agriculture (FFTL4), food (BRFS3, MRFG3, and RNAR3), paper (the sequence made by FIBR3, SUZB5, and KLBN4), and sugar and ethanol (SMT03). There are also stocks belonging to other companies, but they apparently do not form clusters.

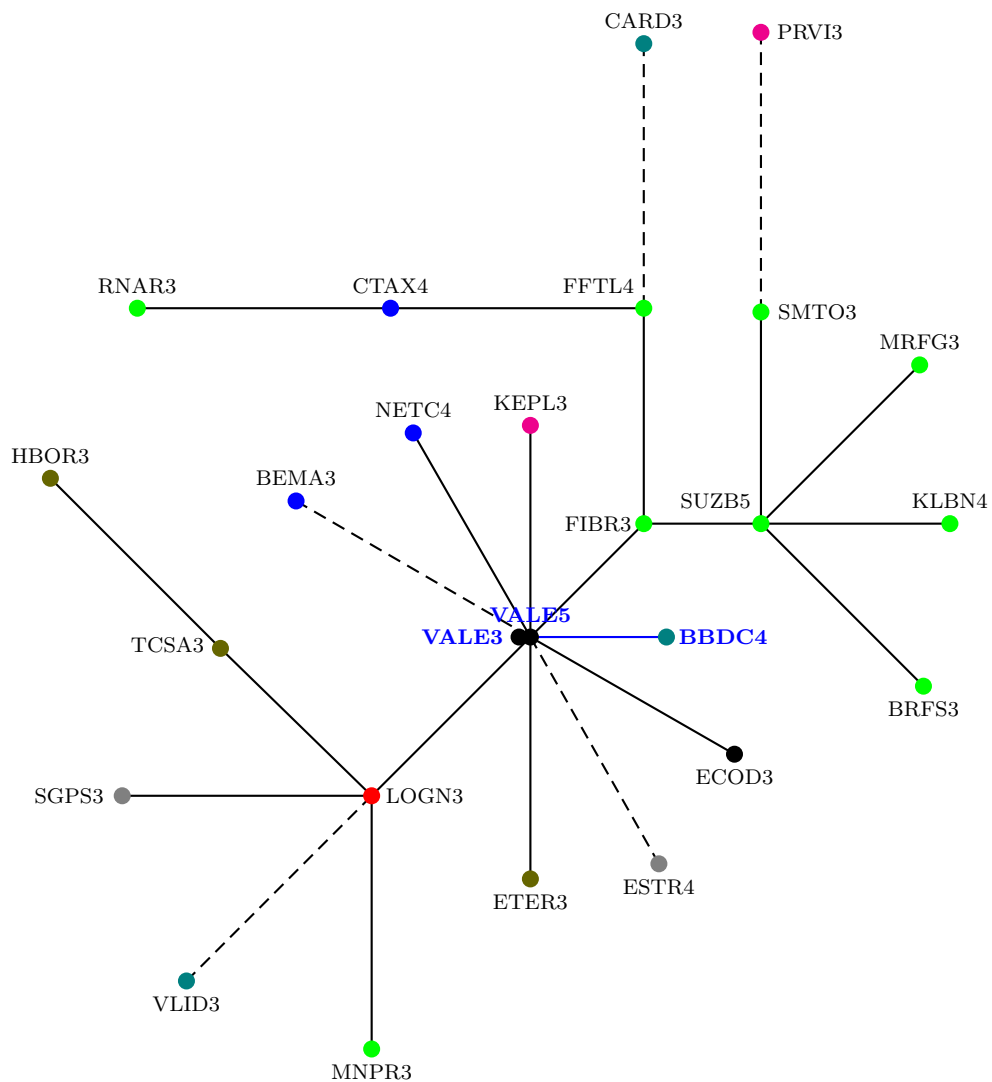


Figure 4: cluster around VALE5.

The cluster in figure 5 is built around BBDC4 (closely connected with BBDC3), which are stocks of Bank Bradesco. It is a denser cluster, comprised in its center by stocks related with banks (ITUB3, ITUB4, and ITSA4, all related with Bank Itau, the second largest in Brazil, SANB3, SANB4, and SANB11, stocks from Bank Santander, BBAS3, stocks of the Bank of Brazil, the largest in Brazil, and BRSR6) or with investment and finance companies (GPIV11, RDCD3), and BVMF3, which are stocks of BM&F-Bovespa itself. To the left, there are two stocks of COSAN (CSAN3 and CZLT11), a food, sugar, and ethanol company. To the top and right, there is another cluster, of cyclic consumer goods: beverages (AMBV3 and AMBV4), tobacco (CRUZ3), pharmaceuticals (DROG3), and sandals (ALPA4). Also to be noticed are the sequences RDCD3-CIEL3, both related with companies that operate and sell credit and debit card terminals, and GOLL4-TAMM4, both related with air transport companies. Another small cluster comprises the stocks TCSL3 and TCSL4, and VIVO4, related with mobile telephony companies. Other stocks scatter around the main network without forming any discernible subnetwork.

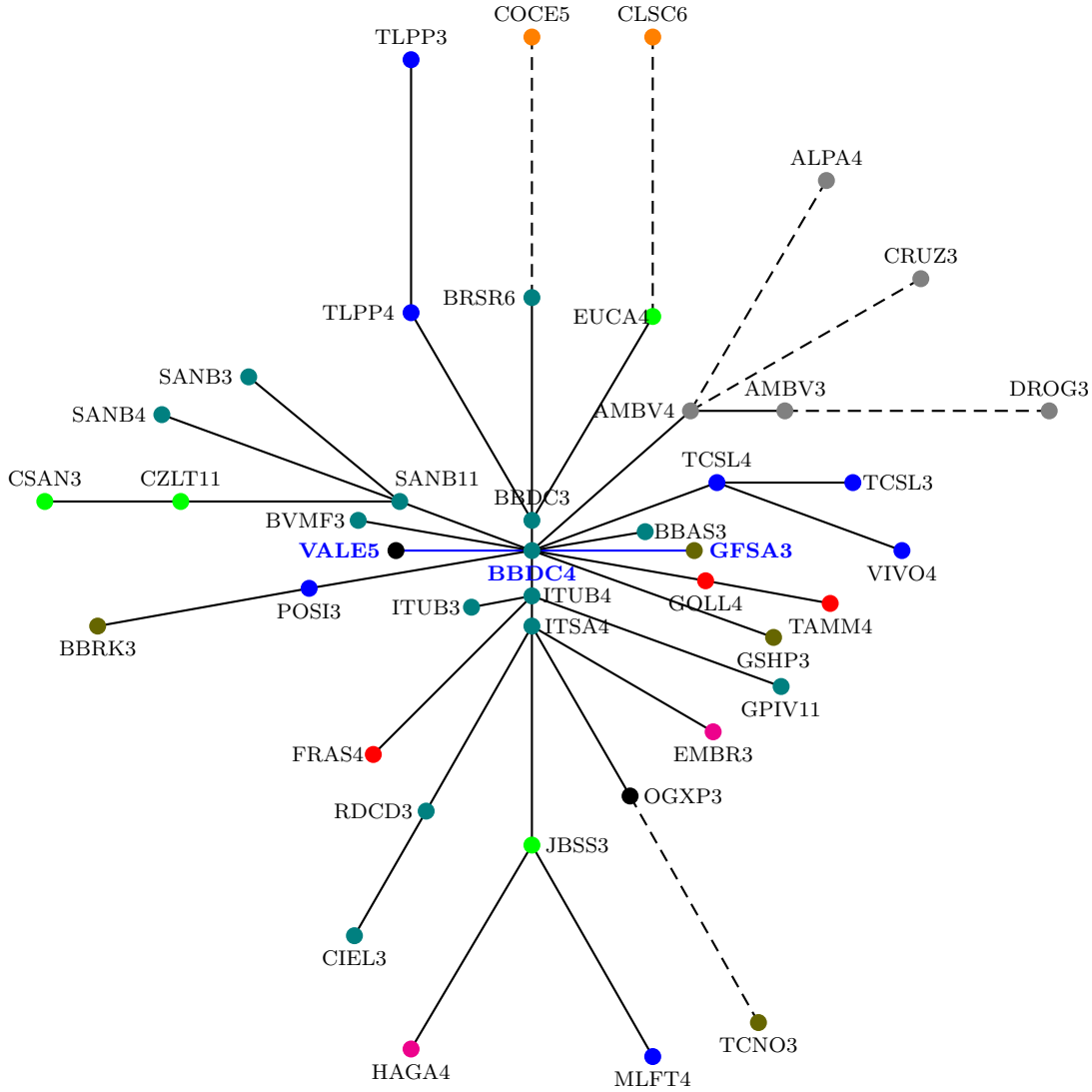


Figure 5: cluster around BBDC4.

The next cluster, represented in figure 6, is centered around a construction and real state company, Gafisa (GFSA3), and is composed mostly by the stocks of other construction and real state companies (PDGR3, CYRE3, CRDE3, EVEN3, INPR3, JFEN3, and LPSB3), and by stocks of consumer goods companies (LAME3 and LAME4, LREN3, BTOW3, NATU3, AMAR3, HGTX3, and MTIG4). There is also a small cluster (lower part the figure) of stocks of electricity distribution companies (GETI3 and GETI4, and TBLE3).

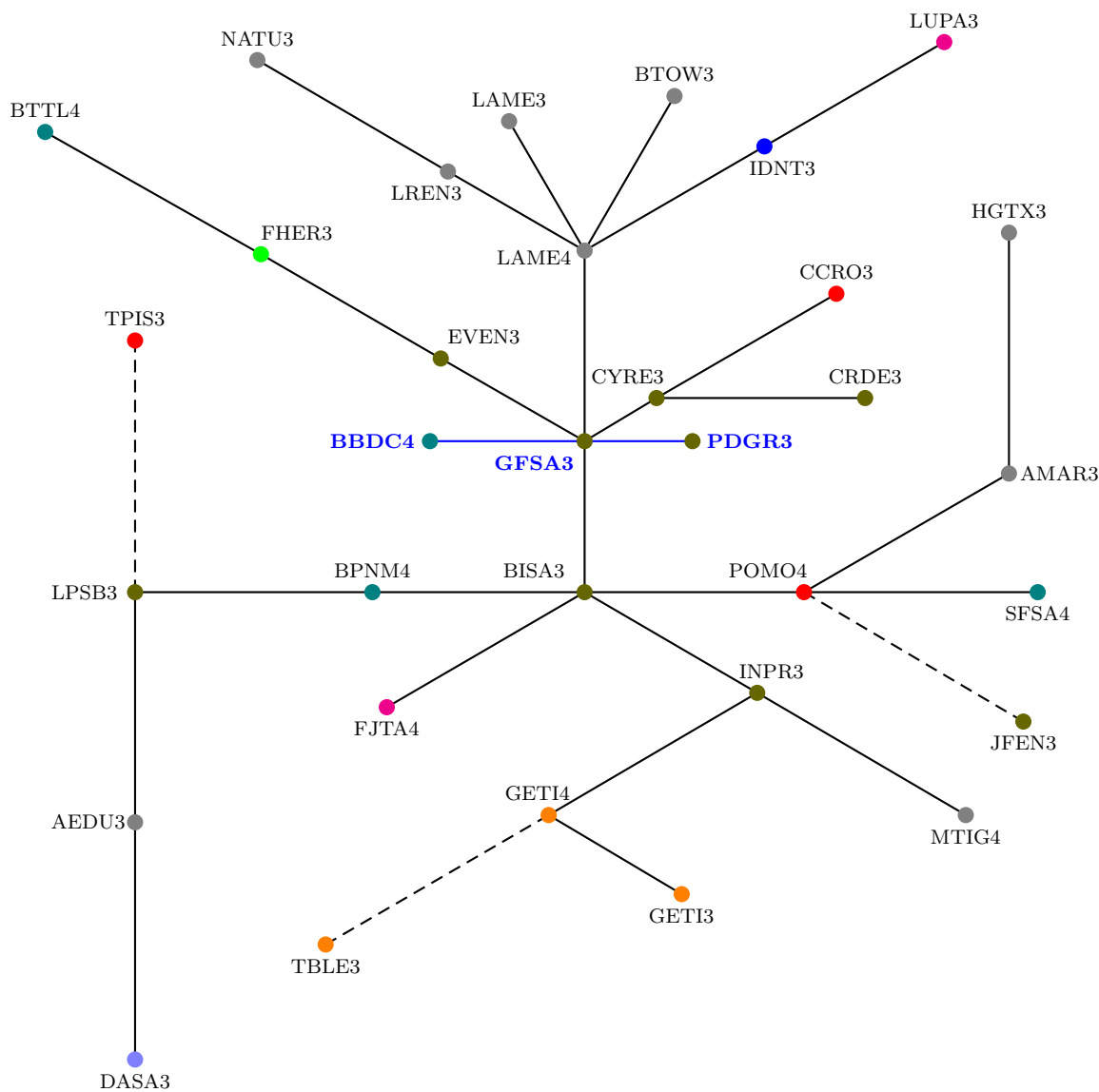


Figure 6: cluster around GFSA3.

The last cluster, figure 7, is also centered around the stock of a construction and real state company, PDG Realty (PDGR3), and is composed mostly of stocks of other construction and real state companies (GFSA3, EZTC3, RSID3, MRVE3, JHSF3, and CCIM3), real state management of shopping centers (BRML3, MULT3, and IGTA3), and of building materials (DTEX3). Other stocks belong to electricity, logistics and transportation, consumer goods and stocks of other types of companies.

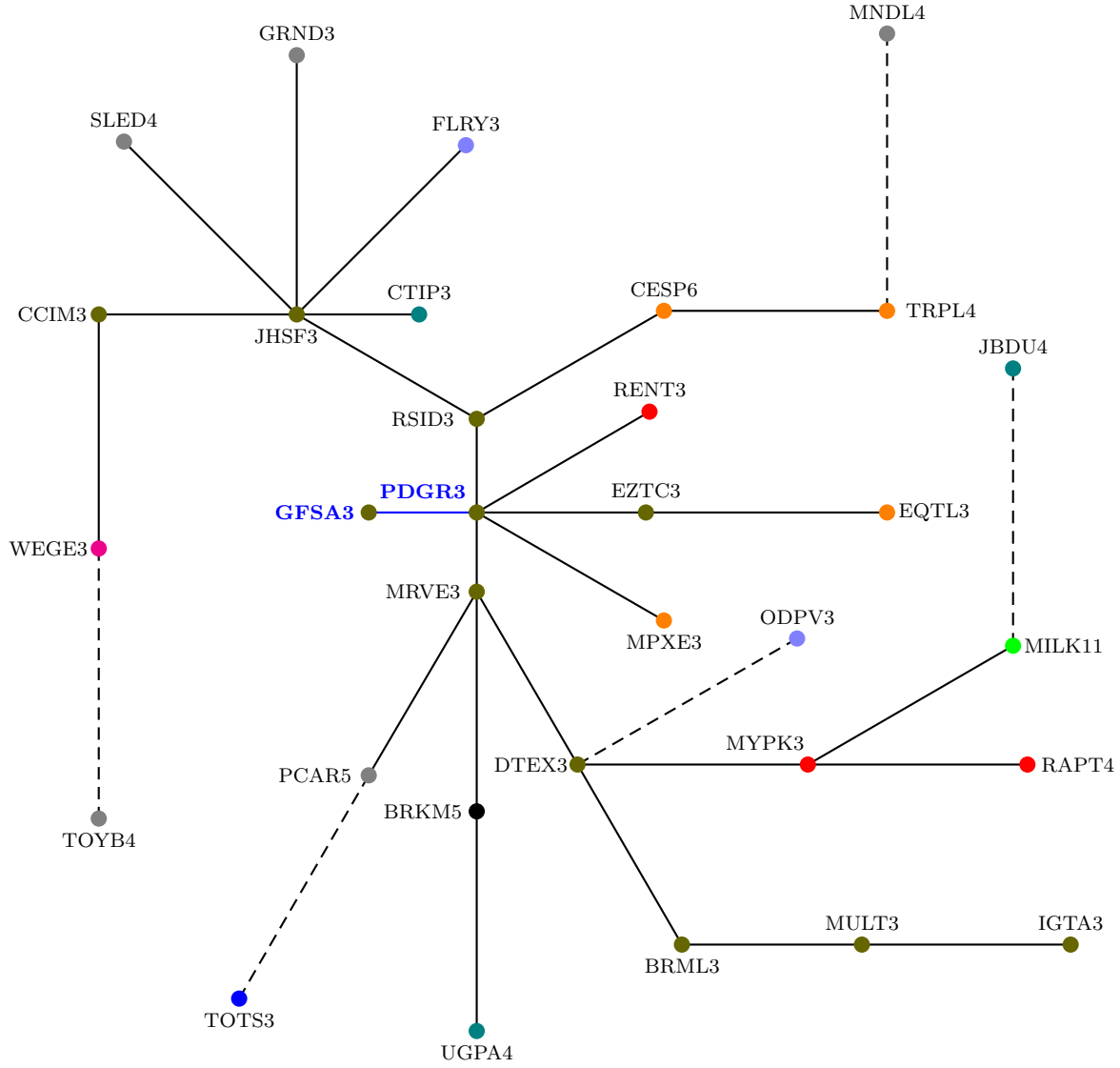


Figure 7: cluster around PDGR3.

3 Centrality measures

In network theory, the centrality of a node, or how influential the vertex is in the network, is an important measurement which is handled in a number of different ways. In what follows, I perform an analysis of the centrality of vertices in the network depicted by the MST of the last section according to five different definitions of centrality. I also do some analysis of the frequency distribution of each centrality measure in the network and which are the stocks that are more central according to each definition.

Node degree

Next, I analyze the node degree of the stocks in the network. The node degree of a node (stock in our case) is the number of connections it has in the network. Most of the stocks have low node degree, and some of them have a large number associated with them. The latter are called hubs, and are generally nodes that are

more important in events that can change the network. Table 1 and figure 8 show the node degree frequency distribution of the network of the BM&F-Bovespa 2010.

Node degree	Frequency
1	105
2	45
3	18
4	7
5	8
6	3
7	0
8	1
9	0
10	1
11	0
12	1
13	1

Table 1: node degree distribution.

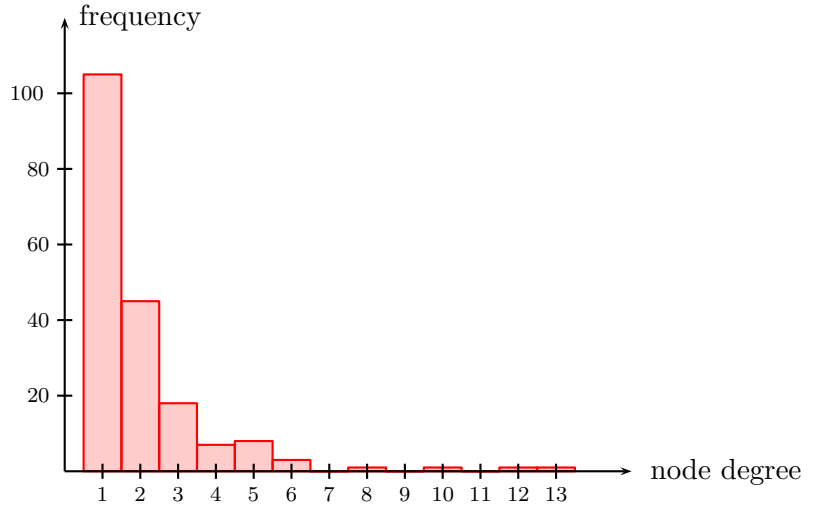


Figure 8: node degree frequency distribution.

The stocks with highest node degree are BRAP4 (node degree 13), BBDC4 (node degree 12), VALE5 (node degree 10), and VALE3 (node degree 8).

Node strength

The strength of a node is the sum of the correlations of the node with all other nodes to which it is connected. If C is the matrix that stores the correlations between nodes that are linked in the minimum spanning tree, then the node strength is given by

$$N_s^k = \sum_{i=1}^n C_{ik} + \sum_{j=1}^n C_{kj} , \quad (3)$$

where C_{ij} is an element of matrix C .

In our network, the vertices with highest node strength are BBDC4 ($N_s = 6.98$), BRAP4 ($N_s = 5.37$), VALE5 ($N_s = 4.72$), and VALE3 ($N_s = 4.50$). The node strength frequency distribution is shown in figure 9.

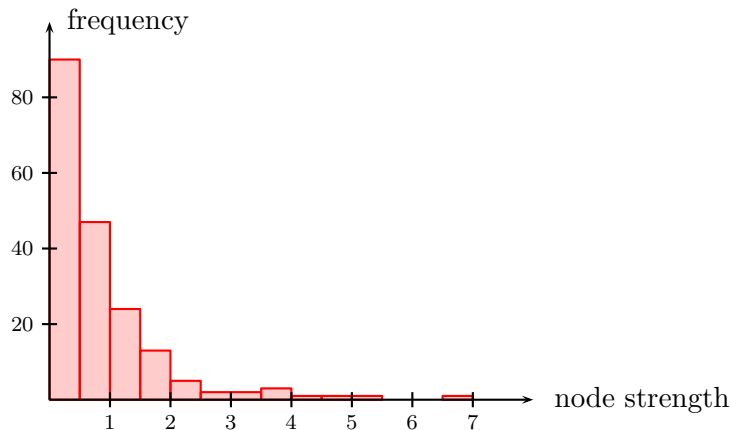


Figure 9: node strength frequency distribution.

Eigenvector centrality

Node degree may be seen as a bad representation of how important a node is, for it doesn't take into account how important the neighbors of a node may be. As an example, one node may have a low degree, but it may be connected with other nodes with very high degree, so it is, in some way, influent. A measure that takes

into account the degree of neighbouring nodes when calculating the importance of a node is called eigenvector centrality. In order to define it properly, one must first define an *adjacency matrix*, A , whose elements a_{ij} are 1 if there is a connection between nodes i and j and zero otherwise. If one now considers the eigenvectors of the adjacency matrix, and choosing its largest value, one then may define the eigenvector with largest eigenvalue by the equation

$$AX = \lambda X , \quad (4)$$

where X is the eigenvector with the largest eigenvalue λ . The eigenvector centrality of a node i is then defined as the i th element of eigenvalue X :

$$E_c = x_i , \quad (5)$$

where x_i is the element of X in row i .

In the present network, the vertices with highest eigenvector centralities are BBDC4 ($E_c = 0.502$), VALE5 ($E_c = 0.373$), VALE3 ($E_c = 0.294$), and BRAP4 ($E_c = 0.257$). The eigenvector centrality frequency distribution is shown in figure 10.

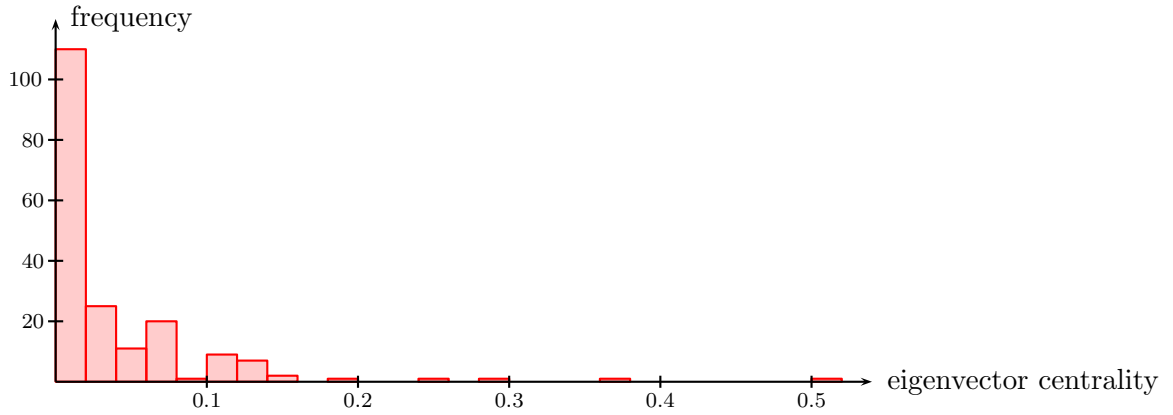


Figure 10: eigenvector centrality frequency distribution.

Betweenness

The betweenness centrality measures how much a node lies on the paths between other vertices. It is an important measure of how much a node is important as an intermediate between other nodes. It may be defined as

$$B_c^k = \sum_{i,j=1}^n \frac{n_{ij}^k}{m_{ij}} , \quad (6)$$

where n_{ij} is the number of shortest paths (geodesic paths) between nodes i and j that pass through node k and m_{ij} is the total number of shortest paths between nodes i and j . Our network is fully connected, so we need not worry about m_{ij} being zero. The betweenness centrality frequency distribution is shown in figure 11.

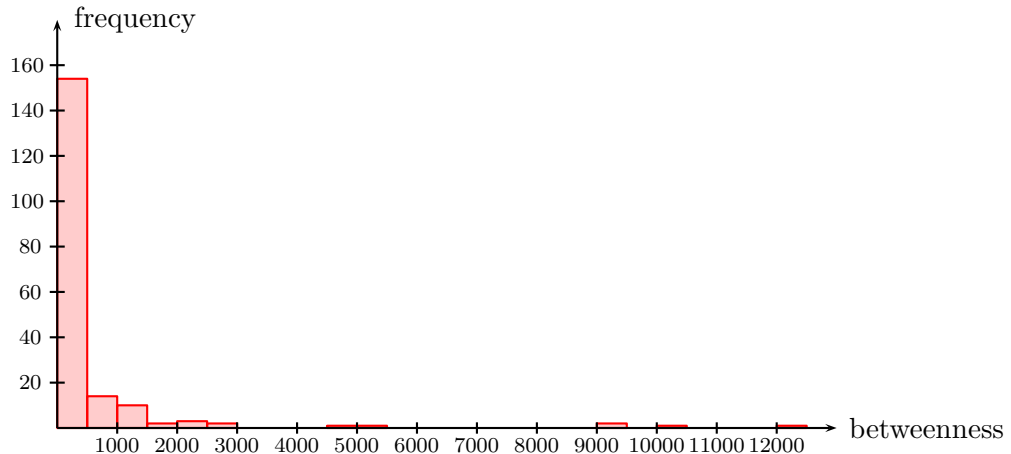


Figure 11: betweenness centrality frequency distribution.

The vertices with highest betweenness centrality are BBDC4 ($B_c = 12010$), VALE5 ($B_c = 10490$), VALE3 ($B_c = 9197$), and GFSA3 ($B_c = 9010$). There are 106 vertices with zero betweenness centrality, what means that no shortest path between any two vertices in the network pass through those vertices.

Closeness

Another measure of centrality is the closeness centrality, which measures the average distance between one node and all the others. It is defined as the measure of the mean geodesic distance for a given node i , which is given by

$$\ell_i = \frac{1}{n} \sum_{j=1}^n d_{ij} , \quad (7)$$

where n is the number of vertices and d_{ij} is a geodesic (minimum path) distance from node i to node j . This measure is small for highly connected vertices and large for distant or poorly connected ones. In order to obtain a measure that is large for highly connected nodes and small for poorly connected ones, one then defines the inverse closeness centrality of node i as

$$C_c^i = \frac{1}{\ell_i} . \quad (8)$$

The vertices with highest inverse closeness centrality are BBDC4 ($C_c = 0.00107$), VALE5 ($C_c = 0.00106$), GFSA3 ($C_c = 0.00100$), and VALE3 ($C_c = 0.00099$). The inverse closeness centrality frequency distribution is shown in figure 12.

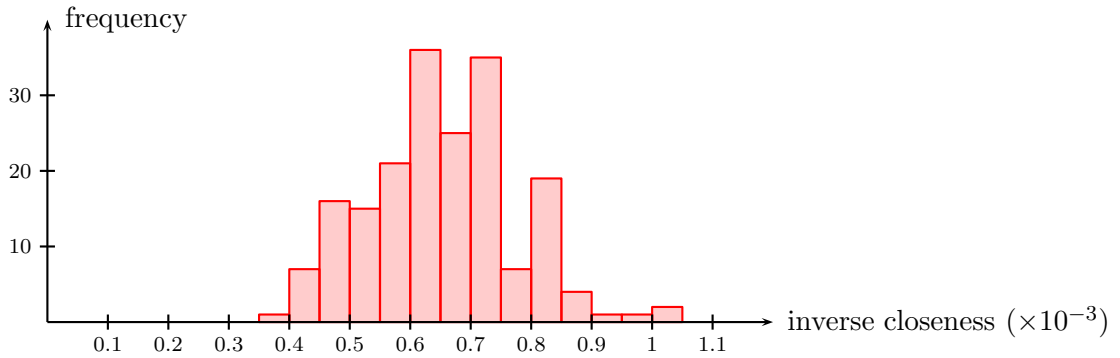


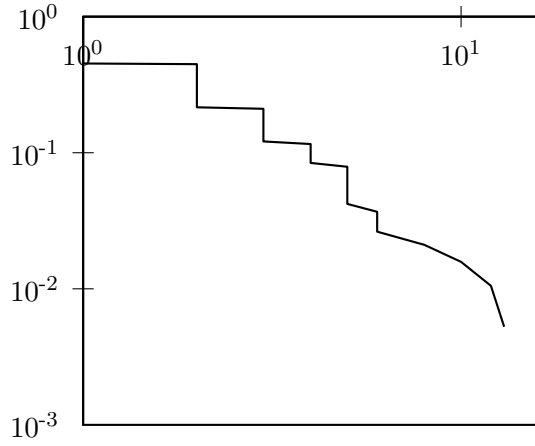
Figure 12: inverse closeness centrality frequency distribution.

4 Cumulative distribution function

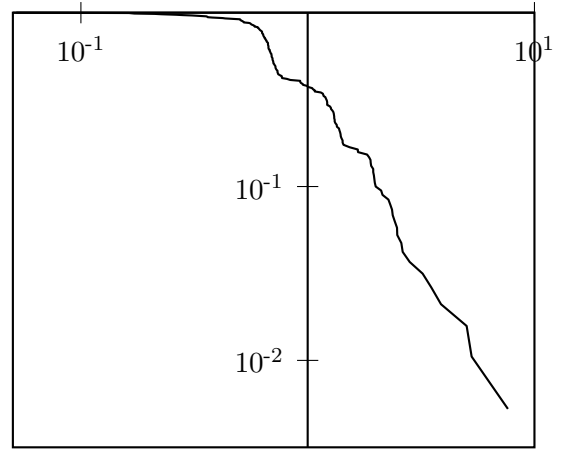
Note that all frequency distributions, with the exception of the one for inverse closeness centrality (the same happens for the frequency distribution of the closeness centrality), are exponentially decreasing. One say that those frequency distributions follow a power law of the type

$$p_k = ck^{-\alpha} , \quad (9)$$

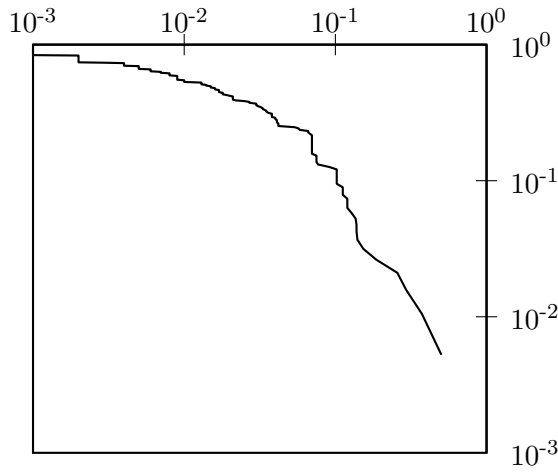
where p_k is the frequency distribution for the value k , and c and α are constants. This is a characteristic of a diversity of complex systems, and it happens in the study of earthquakes, the world wide web, networks of scientific citations, of film actors, of social interactions, protein interactions, and many other topics [37]. Networks whose centrality measures follow this type of distribution are often called scale-free networks, and this behavior can best be visualized if one plots a graph of the cumulative frequency distribution of a centrality in terms of the centrality values, both in logarithmic representation. Figure 13 shows the logarithm of the cumulative frequency distributions as functions of the logarithm of the five types of centrality measures we used in the last section.



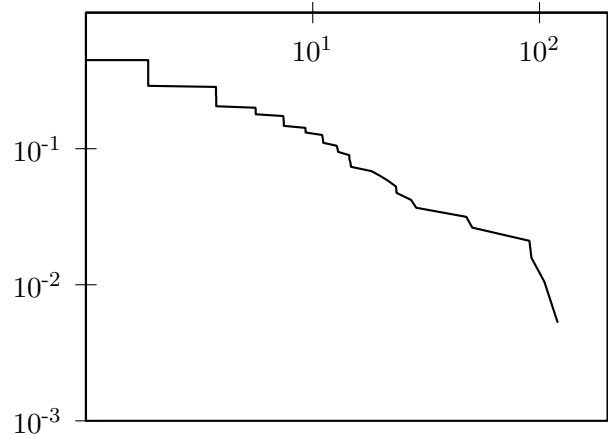
a) Node degree x cumulative distribution



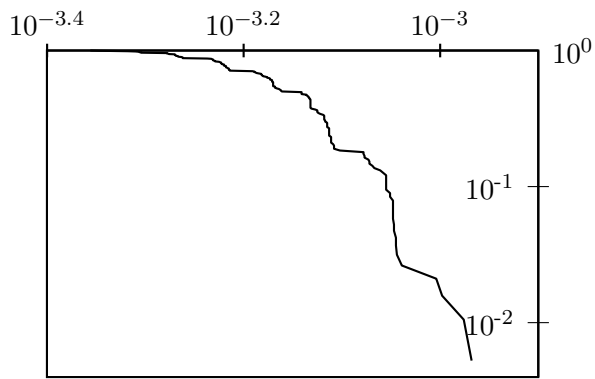
b) Node strength x cumulative distribution



c) Eigenvector centrality x cumulative distribution



d) betweenness centrality x cumulative distribution



e) Inverse closeness x cumulative distribution

Figure 13: log-log plots of the cumulative frequency distributions as functions of centrality measures.

Were a network a pure scale-free one, then the graphic of the logarithm of the cumulative frequency distribution as a function of the logarithm of a centrality measure would be a straight line. One may notice that this behavior is approximately followed by the intermediate levels of all centrality measures, except for inverse closeness centrality. The exactly same behavior is seen in, as examples, the world wide web or networks of citations.

5 Asset Graphs

A two dimensional representation of non-overlapping connections between vertices, such as the minimum spanning tree, has some misleading features. Two nodes that are actually close to each other may appear far or unconnected, and one node that is just very weakly connected may establish a connection in the MST at random. A three dimensional representation is often a better representation, and one can be built using distance as a guide. As an example, one may use an algorithm that minimizes the sum of the squares of the differences between the real distances and the ones portrayed in a three dimensional graph. Other, more advanced techniques, can also be used, such as principal coordinates analysis. By using the latter, one obtains the following three dimensional representation of the stocks of BM&F-Bovespa (figure 14).

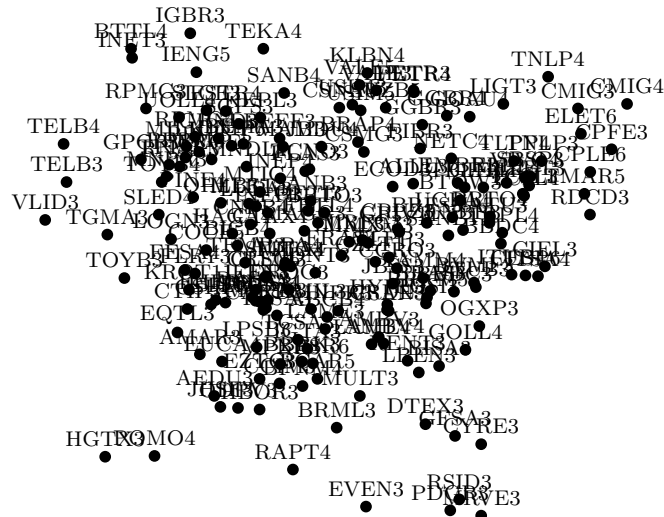


Figure 14: three dimensional representation of stocks of the BM&F-Bovespa.

The picture is not very enlightening when printed on paper, but offers a good three-dimensional view when viewed on a computer. In particular, it may be used in order to study the distributions of stocks of the same type of companies, what is done in Appendix B. Here, we shall use the three dimensional representation in order to study asset graphs, which are networks built on the distance matrix by establishing thresholds under which connections are considered. As an example, one may build a network of those nodes whose distances with the others are below or equal to $T = 0.5$. This will probably exclude many of the connections and some of the nodes of the original network. In what follows, I perform an analysis of some asset graphs obtained by selecting thresholds that go from 0.1 to 0.7, which is the limit at which random noise starts to become absolute.

At $T = 0.1$ (figure 15), the only connections are those between BBDC3 and BBDC4, both stocks from Bradesco (banking), GGBR3, GGBR4, and GOAU4 from Gerdau (metalurgy), ITUB4 and ITSA4 from Itau (banking), PETR3 and PETR4 from Petrobras (petroleum and gas), and between VALE3 and VALE5, both stocks from Vale (mining).

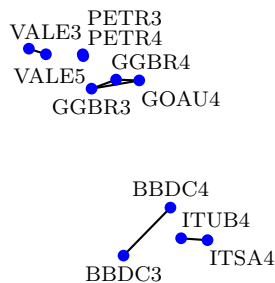


Figure 15: connections bellow threshold $T = 0.1$.

At $T = 0.2$ (figure 16), the earlier connections are joined by a small cluster of stocks RPMG3 and RPMG4, of the Refinaria de Petróleo Mangueiras (petroleum refinement), and by another small cluster formed by USIM3 and USIM4, of Usiminas (mining). Connections are established between BRAP4, Bradespar, which is an investment branch of Bank Bradesco which has 17% of the control of Vale, and VALE3 and VALE5. A cluster is formed with stocks from Bank Bradesco and Bank Itau, now with the joining of ITUB3.

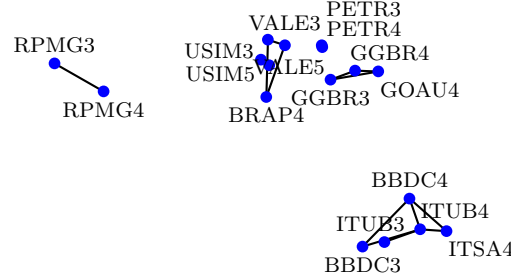


Figure 16: connections bellow threshold $T = 0.2$.

At $T = 0.3$ (figure 17), the network is joined by the pairs TELB3-TELB4 of Telebras (telecommunications), INEP3-INEP4 of Inepar (construction), TNLP3-TNLP4 of Telemar (telecommunications), CMIG3-CMIG4 of CEMIG (electricity), and AMBV3-AMBV4 of Ambev (beverages). The network formed by the stocks of Gerdau now joins with the stocks of Usiminas, and with the newcomer CSNA3 of Companhia Siderúrgica Nacional (metalurgy). The new network represents a cluster of stocks of companies working in mining and meatlurgy, close to but still not connected with Vale. A new newtork is formed by the stocks GFSA3 of Gafisa, CYRE3 of Cyrela, RSID3 of Rossi Residencial, PDGR3 of PDG Realty, and MRVE3 of MRV Engenharia e Participações, all of them in the construction business.

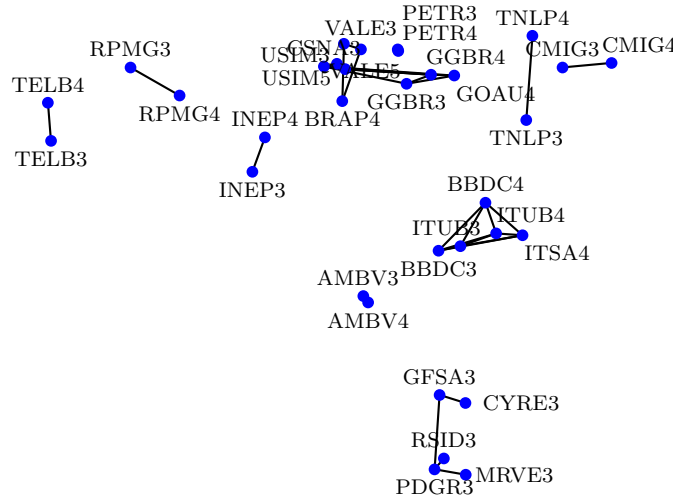


Figure 17: connections bellow threshold $T = 0.3$.

For $T = 0.4$ (figure 18, with the new stocks in the network in red, for better visualization), the mining and metalurgy network joins with the financial network, and the construction network becomes denser, with more connections established between its nodes. We also have the newcomer isolated pairs TCSL3-TCSL4 of TIM (telecommunications), CIEL3 of Cielo and RDCD3 of Redecard, both operating in the business of electronic cards, SUZB5 of Suzano Papel e Celulose and FIBR3 of Fibria Celulose, both operating in the paper production market, and GOLL4 of Gol and TAMM4 of TAM, both pertaining to airlines. The stocks of two banks, SANB11 of Bank Santander and BBAS3 of the Bank of Brazil connect with the financial network via the stocks of Bradesco, and the stocks CZLT11 of Cosan (sugar and ethanol) connect with CSNA3 (metalurgy).

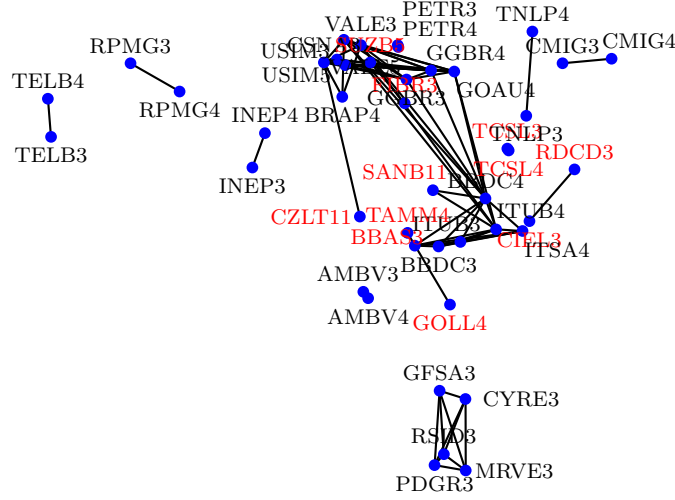


Figure 18: connections bellow threshold $T = 0.4$.

For $T = 0.5$ (figure 19, with the new stocks in the network highlighted in red), all three major networks, mining and metalurgy, financial, and construction, are now joined, with a central role played by the financial network. BISA3 of Brookfield Incorporações, EVEN3 of Even Construtora e Incorporadora, EZTC3 of EZETEC, all of the building industry, join the building companies network. BRML3 of BR Malls and MULT3 of Multiplan, both of the real state business (shopping centers), form a pair close to but not connected with the building industry network.

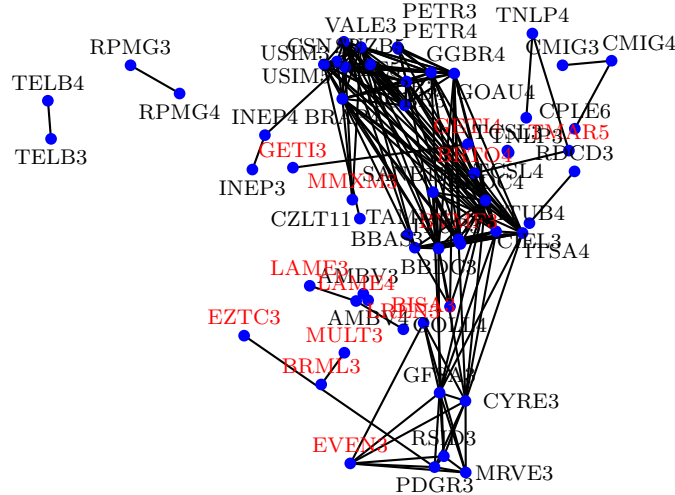


Figure 19: connections bellow threshold $T = 0.5$.

BRTO4 of Brasil Telecom and TMAR5 of Telemar Norte Leste, both of telecommunications companies, join TNLP4 of Telemar in order to form a telecommunications network. GETI3 and GETI4, both of AES Tietê (electricity) form an isolated pair. LAME3 and LAME4 of Lojas Americanas and LREN3 of Lojas Renner, both of the retailing business, form another separate cluster. MMXM3 of MMX Mineração e Metálicos join Vale and Bradespar in the mining and metalurgy network. The pair of stocks of Inepar (construction) connects with Vale (mining), the pair of Telemar connects with Telemar Norte Leste, the stocks of Gol (airliner) connect with the ones of Bradesco (banking) and Gerdau (metalurgy). BVMF3, which are stocks of the BM&F-Bovespa (financial), connect with the ones of Bradesco, Bradespar, Itau (all in the banking and financial sectors), Cyrella (building), and Gerdau (metalurgy). Petrobras (petroleum and gas) connects with Bradesco (banking) and Gerdau (metalurgy).

For $T = 0.6$ (figure 20, with the new stocks in the network highlighted in red), and $T = 0.7$, connections between sectors become more frequent, and already existing clusters become denser. More stocks take part of the complete network now, even those that have weaker correlations with other stocks. From $T = 0.8$ onwards, noise starts to takes over, and new connections are not reliable.

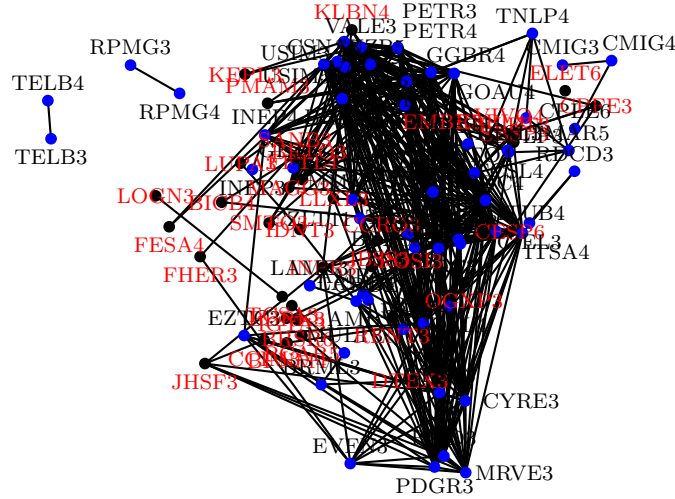


Figure 20: connections bellow threshold $T = 0.5$.

6 Centrality measures for the asset graph

I shall now analyze some of the centrality measures seen in sections 3 and 4 but now applying them to the network obtained by fixing a threshold $T = 0.7$ (just bellow the noisy region) and considering only those distances that are bellow it. I begin by analyzing the node degree. Table 2 and figure 21 show the node degree distribution.

Node degree	Frequency
0	22
10	47
20	22
30	10
40	14
50	13
60	16
70	8
80	9
90	6
100	4
110	3
120	6
130	3
140	2
150	3
160	2

Table 2: node degree distribution for the asset graph.

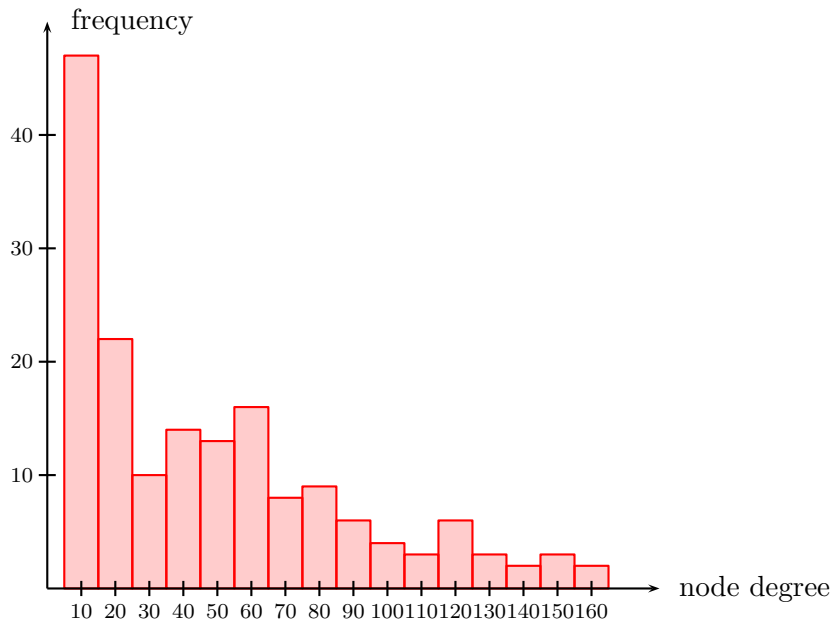


Figure 21: node degree frequency distribution.

The highest node degrees belong to BBDC4 and ITUB4 (node degree 103), BRAP4 (node degree 101), ITSA4 (node degree 95), BBDC3, and VALE5 (node degree 95). Figure 22 represents the acumulated frequency distribution of the node degree plotted against the node degree.

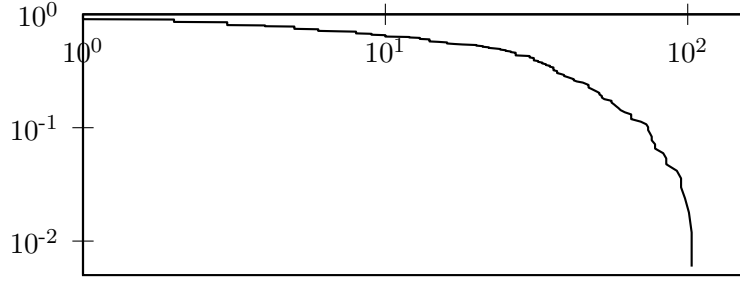


Figure 22: node degree x cumulative distribution

Next, calculating the node strenght, one can see that the nodes with highest values are BBDC4 ($N_s = 86.87$), ITUB4 ($N_s = 85.05$), BRAP4 ($N_s = 81.22$), and ITSA4 ($N_s = 79.93$). The probability distribution function is shown in the figure bellow, together with the log-log plot of the cumulative distribution as a function of the node strenght.

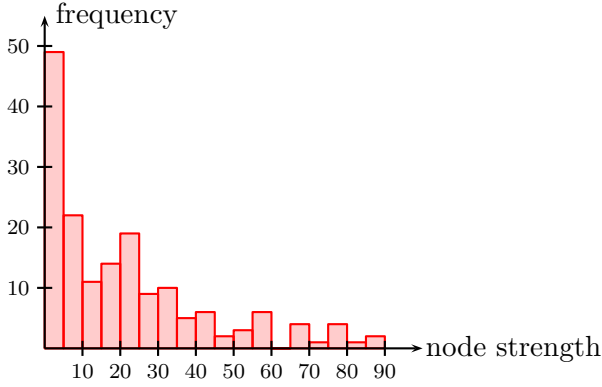


Figure 23: node strength frequency distribution.

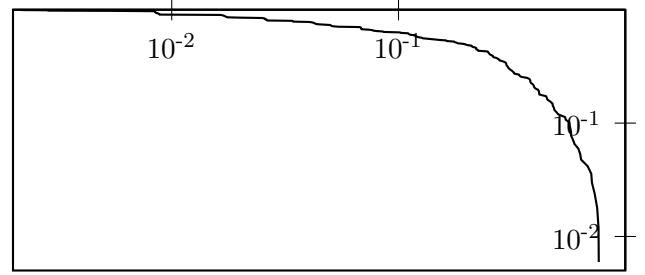


Figure 24: node strength x cumulative distribution.

The eigenvector centrality, given by equation (4), measures the centrality of a node by giving weights to the nodes that are immediately connected with it. The highest values for eigenvector centrality are those of BBDC4 and ITUB4 ($E_c = 0.172$), ITSA4 ($E_c = 0.167$), BBDC3 and BRAP4 ($E_c = 0.163$), and VALE5 ($E_c = 0.160$). The probability distribution function is shown in figure 25, together with the log-log plot of the cumulative distribution as a function of the eigenvector centrality (figure 26).

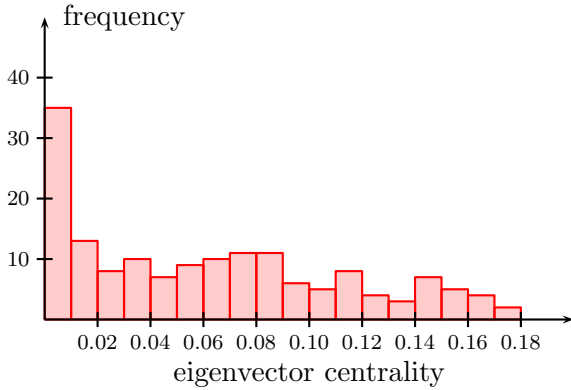


Figure 25: eigenvector centrality frequency distribution.

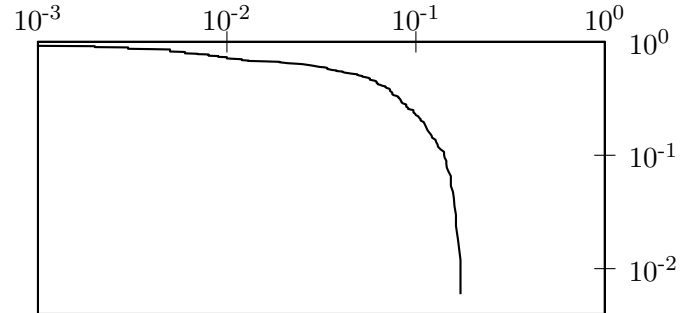


Figure 26: eigenvector centrality x cumulative distribution

If we now consider betweenness as a measure of centrality, then the highest values occur for BRAP4 ($B_c = 950$), BBDC4 ($B_c = 722$), ITUB4 ($B_c = 701$), and VALE5 ($B_c = 629$). The probability distribution function is shown in figure 27, together with the log-log plot of the cumulative distribution as a function of the betweenness centrality (figure 28).

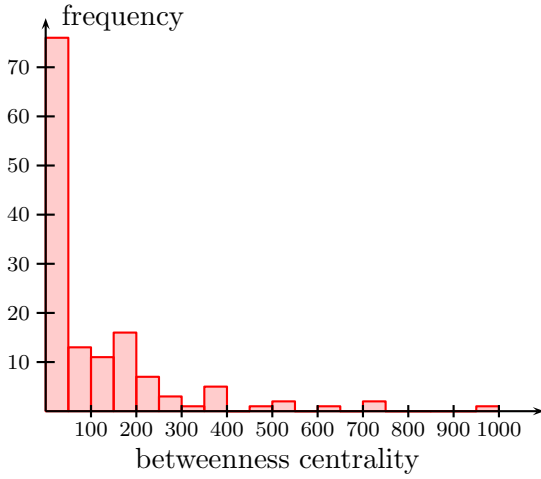
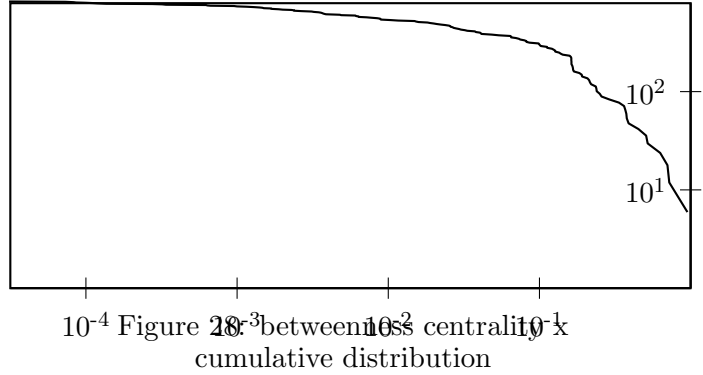


Figure 27: betweenness centrality frequency distribution.



The last centrality measure I shall analyze is inverse closeness centrality, for which the highest values occur for BBDC4 and ITUB4 ($C_c = 1.866 \times 10^{-4}$), BRAP4 and ITSA4 ($C_c = 1.864 \times 10^{-4}$), BBDC3 ($C_c = 1.863 \times 10^{-4}$), and VALE5 ($C_c = 1.862 \times 10^{-4}$). The probability distribution function is shown in figure 29, together with the log-log plot of the cumulative distribution as a function of the inverse closeness centrality (figure 30).

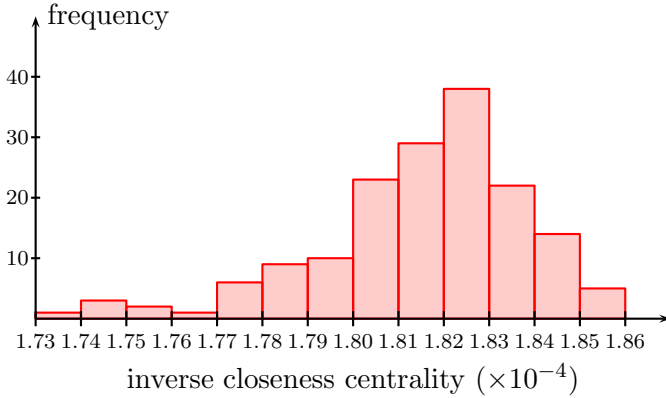


Figure 29: inverse closeness centrality frequency distribution.

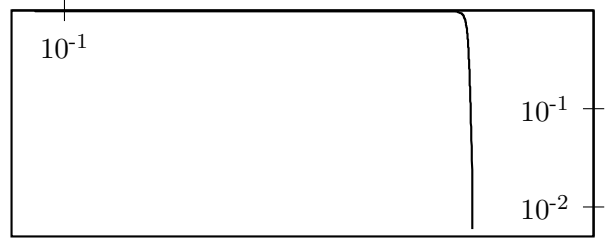


Figure 30: inverse closeness centrality x cumulative distribution

One thing to be noticed is that the frequency distributions of the centrality measures of the asset graph are more distant from what would be expected from a scale-free network. This may be due to the great amount of noise that comes with the choice of threshold. For higher threshold values, this difference increases, and for lower threshold values, it decreases.

7 K-shell decomposition

Another important centrality measure which can be applied only to the asset graph (and not to the MST) is k-shell decomposition, which consists on classifying a vertex according to the connections it makes, and also considers if it is in a region of the network that is also highly connected. This decomposition is frequently used in the study of the propagation of diseases and of information, and also shed some light on financial networks. It consists of considering all nodes with degree 1, assigning to them $k = 1$, and stripping the network from them. Then, one looks at the remaining vertices with degree 2 or less, and assigns to them $k = 2$. Repeating the procedure, all vertices with $k = 2$ are removed, and one then looks for vertices with degree 3 or less. The process goes on until all nodes are removed. What one then obtains are shells of vertices that increase in importance as k goes larger. For our asset graph network for $T = 0.7$, there are 30 shells, and the stocks that belong to it are represented in figure 31.

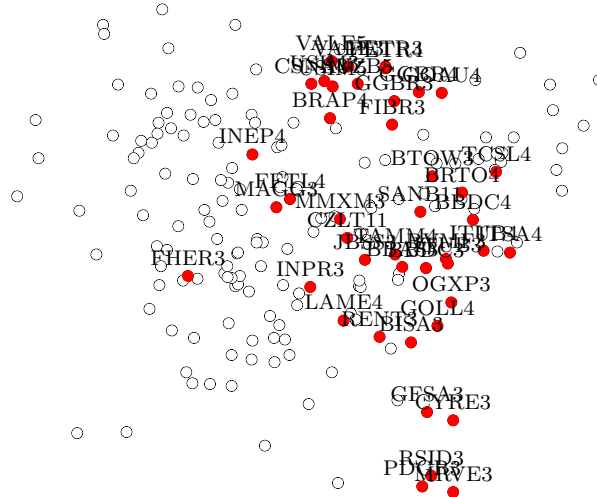


Figure 31: core ($k = 30$) of the asset graph.

There is a strong dependence of k -shell position and node degree, as can be seen in figure 32. With the exception of the innermost shell, there is an almost linear relation between the two measures.

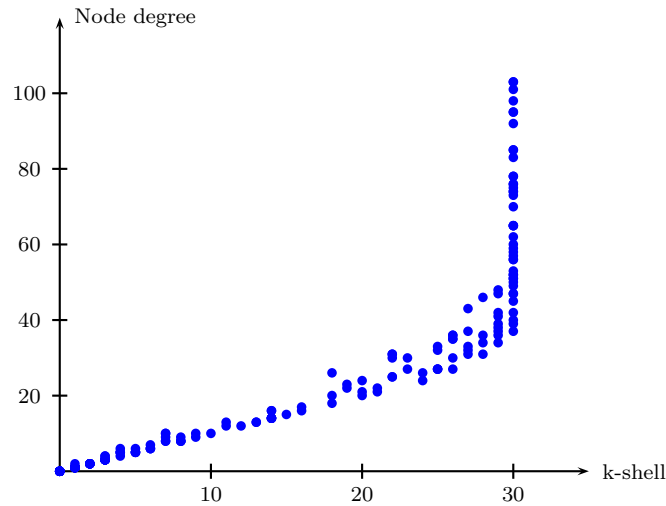


Figure 32: node degree as a function of k -shell number.

8 Conclusion

Using the concepts of minimum spanning trees and asset graphs, both based on the correlation matrix of log-returns of the BM&F-Bovespa for the year 2010, maps of that stock exchange were built, and the network structure so obtained was examined. One could see that the BM&F-Bovespa has a clustered structure roughly based on economic activity sectors, and gravitates around some key stocks. The results may be used in order to obtain a better knowledge of the structure of this particular stock exchange, and maybe devise diversification strategies for portfolios of stocks belonging to it. There was also the opportunity to compare two different representations of the same complex system, each one with its benefits and maladies.

Acknowledgements

I thank for the support of this work by a grant from Insper, Instituto de Ensino e Pesquisa. I am also grateful to Leonidas Sandoval Neto (my father and Economist), who helped me clarify some details, and to Gustavo Curi Amarante, who collected the data. This article was written using \LaTeX , all figures were made

using PSTricks, and the calculations were made using Matlab, Ucinet and Excel. All data are freely available upon request on leonidassj@insper.edu.br.

A Stocks, codes, and sectors

Here I display, in alphabetical order, the stocks that are being used in the present work, together with the companies they represent and the sectors those belong to. They are not all the stocks that are negotiated in the BM&F-Bovespa, for only the ones that were negotiated every day the stock market opened are being considered. So, all the stocks have high liquidity and there is no missing data in the time series being used.

Code	Company	Sector
<i>ABCB4</i>	Banco ABC	Bancário
<i>AEDU3</i>	Anhanguera Educacional	Educação
<i>ALLL3</i>	América Latina Logística	Logística e transporte - trens
<i>ALPA4</i>	Alpargatas	Tecidos, vestuário e calçados
<i>AMAR3</i>	Lojas Marisa	Tecidos, vestuário e calçados
<i>AMBV3</i>	AMBEV	Consumo - bebidas
<i>AMBV4</i>	AMBEV	Consumo - bebidas
<i>AMIL3</i>	Amil Participações	Saúde
<i>BBAS3</i>	Banco do Brasil	Bancário
<i>BBDC3</i>	Bradesco	Bancário
<i>BBDC4</i>	Bradesco	Bancário
<i>BBRK3</i>	Brasil Brokers	Construção, Materiais e Engenharia
<i>BEEF3</i>	Minerva	Alimentos
<i>BEMA3</i>	Bematech	Tecnologia, internet e call-centers
<i>BICB4</i>	Bicbanco	Bancário
<i>BISA3</i>	Brookfield Incorporações	Construção, Materiais e Engenharia
<i>BVMF3</i>	BM&F-BOVESPA	Financeiro
<i>BPNM4</i>	Banco Panamericano	Bancário
<i>BRAP4</i>	Bradespar	Investimentos
<i>BRFS3</i>	Brasil Foods	Alimentos
<i>BRKM5</i>	Braskem	Petroquímica
<i>BRML3</i>	BR Malls	Imóveis
<i>BRSR6</i>	BANRISUL	Bancário
<i>BRTO3</i>	Brasil Telecom	Telefonia fixa
<i>BRTO4</i>	Brasil Telecom	Telefonia fixa
<i>BTOW3</i>	B2W - Companhia Global do Varejo	Consumo - eletrônicos
<i>BTTL4</i>	Battistella Adm	Investimentos
<i>BVMF3</i>	BM&F-BOVESPA	Financeiro
<i>CARD3</i>	CSU Cardsystem	Serviços diversos
<i>CCIM3</i>	Camargo Correa	Construção, Materiais e Engenharia
<i>CCRO3</i>	CCR	Logística e transporte
<i>CESP6</i>	CESP	Energia elétrica
<i>CIEL3</i>	Cielo	Financeiro
<i>CLSC6</i>	Celesc	Energia elétrica
<i>CMIG3</i>	CEMIG	Energia elétrica
<i>CMIG4</i>	CEMIG	Energia elétrica
<i>CNFB4</i>	Confab	Siderurgia
<i>COCE5</i>	Coelce	Energia elétrica
<i>CPFE3</i>	CPFL	Energia elétrica
<i>CPLE6</i>	COPEL	Energia elétrica
<i>CRDE3</i>	CR2 Empreendimentos Imobiliários	Construção, Materiais e Engenharia
<i>CREM3</i>	Cremer	Saúde
<i>CRUZ3</i>	Souza Cruz	Tobacco
<i>CSAN3</i>	COSAN	Açúcar e álcool
<i>CSMG3</i>	COPASA	Saneamento
<i>CSNA3</i>	Cia Siderúrgica Nacional	Siderurgia

Code	Company	Sector
<i>CTAX4</i>	Contax	Tecnologia, internet e call-centers
<i>CTIP3</i>	CETIP	Financeiro
<i>CYRE3</i>	Cyrela Brazil	Construção, Materiais e Engenharia
<i>CZLT11</i>	COSAN	Alimentos
<i>DASA3</i>	DASA - Diagnósticos da América	Saúde
<i>DROG3</i>	Drogasil	Consumo - farmacêutico
<i>DTEX3</i>	Duratex	Construção, Materiais e Engenharia
<i>ECOD3</i>	Brasil Ecodiesel	Petróleo, gás e biocombustíveis
<i>ELET3</i>	Eletrobras	Energia elétrica
<i>ELET6</i>	Eletrobras	Energia elétrica
<i>ELPL4</i>	Eletropaulo	Energia elétrica
<i>EMBR3</i>	Embraer	Indústria de aviação
<i>ENBR3</i>	EDP - Energias do Brasil	Energia elétrica
<i>EQTL3</i>	Equatorial	Energia elétrica
<i>ESTR4</i>	Estrela	Consumo - brinquedos
<i>ETER3</i>	Eternit	Construção, Materiais e Engenharia
<i>EUCA4</i>	Eucatex	Madeira e papel
<i>EVEN3</i>	Even Construtora e Incorporadora	Construção, Materiais e Engenharia
<i>EZTC3</i>	EZTEC Empreendimentos e Participações	Construção, Materiais e Engenharia
<i>FESA4</i>	FERBASA	Siderurgia
<i>FFTL4</i>	Valefert	Agronegócios
<i>FHER3</i>	Fertilizantes Heringer	Agronegócios
<i>FIBR3</i>	Fibria Celulose	Papel e celulose
<i>FJTA4</i>	Forjas Taurus	Máquinas e equipamentos
<i>FLRY3</i>	Fleury	Saúde
<i>FRAS4</i>	Fras-Le	Material de transporte
<i>GETI3</i>	AES Tietê	Energia elétrica
<i>GETI4</i>	AES Tietê	Energia elétrica
<i>GFA3</i>	Gafisa	Construção, Materiais e Engenharia
<i>GGBR3</i>	Gerdau	Siderurgia
<i>GGBR4</i>	Gerdau	Siderurgia
<i>GOAU4</i>	Metalúrgica Gerdau	Siderurgia - metalurgia
<i>GOLL4</i>	GOL	Logística e transporte - aviação
<i>GPCP3</i>	GPC Participações	Petroquímica
<i>GPIV11</i>	GP Investments	Investimentos
<i>GRND3</i>	Grendene	Tecidos, vestuário e calçados
<i>GSHP3</i>	General Shopping Brasil	Imóveis
<i>HAGA4</i>	Haga	Materiais de construção
<i>HBOR3</i>	Helbor	Construção, Materiais e Engenharia
<i>HGTX3</i>	Hering	Tecidos, vestuário e calçados
<i>HYPE3</i>	Hypermarcas	Consumo
<i>IDNT3</i>	Ideiasnet	Tecnologia, internet e call-centers
<i>IENG5</i>	INEPAR	Construção, Materiais e Engenharia
<i>IGBR3</i>	IGB Eletrônica	Eletrônica
<i>IGTA3</i>	Iguatemi	Imóveis
<i>INEP3</i>	Inepar Indústria e Construções	Máquinas e equipamentos
<i>INEP4</i>	Inepar Indústria e Construções	Máquinas e equipamentos
<i>INET3</i>	INEPAR Telecomunicações	Telecomunicações
<i>INPR3</i>	Inpar	Construção, Materiais e Engenharia
<i>ITSA4</i>	Itausa	Bancário
<i>ITUB3</i>	Itau Unibanco	Bancário
<i>ITUB4</i>	Itau Unibanco	Bancário
<i>JBDU4</i>	JB Duarte	Investimentos
<i>JBSS3</i>	JBS	Alimentos
<i>JFEN3</i>	João Fortes Engenharia	Construção, Materiais e Engenharia
<i>JHSF3</i>	JHSP Participações	Construção, Materiais e Engenharia
<i>KEPL3</i>	Kepler Weber	Máquinas e equipamentos

Code	Company	Sector
<i>KLBN4</i>	Klabin	Papel e celulose
<i>KROT11</i>	Kroton Educacional	Educação
<i>LAME3</i>	Lojas Americanas	Consumo - comércio
<i>LAME4</i>	Lojas Americanas	Consumo - comércio
<i>LIGT3</i>	Light	Energia elétrica
<i>LLXL3</i>	LLX Logística	Logística e transporte
<i>LOGN3</i>	Log-in Logística Intermodal	Logística e transporte
<i>LPSB3</i>	Lopes Brasil	Construção, Materiais e Engenharia
<i>LREN3</i>	Lojas Renner	Consumo - vestuário
<i>LUPA3</i>	Lupatech	Máquinas e equipamentos
<i>MAGG3</i>	Magnesita Refratários	Siderurgia
<i>MILK11</i>	LAEP Investments	Alimentos
<i>MLFT4</i>	Jereissati Participações	Telefonia fixa
<i>MMXM3</i>	MMX Mineração e Metálicos	Mineração
<i>MNDL4</i>	Mundial	Consumo - eletrodomésticos
<i>MNPR3</i>	Minupar	Alimentos
<i>MPXE3</i>	MPX Energia	Energia elétrica
<i>MRFG3</i>	MARFRIG	Alimentos
<i>MRVE3</i>	MRV Engenharia e Participações	Construção, Materiais e Engenharia
<i>MTIG4</i>	Metalgráfica Iguaçu	Consumo - embalagens
<i>MULT3</i>	Multiplan	Imóveis
<i>MYPK3</i>	IOCHPE Maxion	Material de transporte
<i>NATU3</i>	Natura	Consumo - cosméticos
<i>NETC4</i>	NET	Telecomunicações
<i>ODPV3</i>	Odontoprev	Saúde
<i>OGXP3</i>	OGX Petróleo e Gás	Petróleo, gás e biocombustíveis
<i>OHLB3</i>	Obrascon Huarte Lain Brasil	Logística e transporte
<i>PCAR5</i>	Pão de Açúcar	Consumo - comércio
<i>PDGR3</i>	PDG Realty	Construção, Materiais e Engenharia
<i>PETR3</i>	Petrobras	Petróleo, gás e biocombustíveis
<i>PETR4</i>	Petrobras	Petróleo, gás e biocombustíveis
<i>PFRM3</i>	Proframa	Consumo - farmacêutico
<i>PINE4</i>	PINE	Bancário
<i>PLAS3</i>	Plascar	Material de transporte
<i>PMAM3</i>	Paranapanema	Material de transporte
<i>POMO4</i>	Marcopolo	Material de transporte
<i>POS13</i>	Positivo Informática	Tecnologia, internet e call-centers
<i>PRVI3</i>	Providência	Materiais diversos
<i>PSSA3</i>	Porto Seguro	Seguradora
<i>RAPT4</i>	Randon	Material de transporte
<i>RDCD3</i>	Redecard	Financeiro
<i>RENT3</i>	Localiza Rent a Car	Logística e transporte - aluguel
<i>RNAR3</i>	Renar Macas	Alimentos
<i>RPMG3</i>	Refinaria de Petróleos Manguinhos	Petróleo, gás e biocombustíveis
<i>RPMG4</i>	Refinaria de Petróleos Manguinhos	Petróleo, gás e biocombustíveis
<i>RSID3</i>	Rossi Residencial	Construção, Materiais e Engenharia
<i>SANB11</i>	Banco Santander	Bancário
<i>SANB3</i>	Santander	Bancário
<i>SANB4</i>	Santander	Bancário
<i>SBSP3</i>	Sabesp	Saneamento
<i>SFSA4</i>	Sofisa	Bancário
<i>SGPS3</i>	Spring Global	Tecidos, vestuário e calçados
<i>SLCE3</i>	SLC Agrícola	Agronegócios
<i>SLED4</i>	Saraiva	Consumo - livros
<i>SMT03</i>	São Martinho	Açúcar e álcool
<i>SULA11</i>	Sul América	Seguradora
<i>SUZB5</i>	Suzano Papel e Celulose	Papel e celulose

Code	Company	Sector
<i>TAMM4</i>	TAM	Logística e transporte - aviação
<i>TBLE3</i>	Tractebel	Energia elétrica
<i>TCNO3</i>	Tecnosolo Engenharia	Construção, Materiais e Engenharia
<i>TCSA3</i>	Tecnisa	Construção, Materiais e Engenharia
<i>TCSL3</i>	TIM	Telefonia móvel
<i>TCSL4</i>	TIM	Telefonia móvel
<i>TEKA4</i>	TEKA	Tecelagem
<i>TELB3</i>	Telebras	Telefonia fixa
<i>TELB4</i>	Telebras	Telefonia fixa
<i>TEMP3</i>	Tempo	Saúde
<i>TGMA3</i>	Tegma Gestão Logística	Logística e transporte
<i>TLPP3</i>	Telesp	Telefonia fixa
<i>TLPP4</i>	Telesp	Telefonia fixa
<i>TMAR5</i>	Telemar Norte Leste	Telefonia fixa
<i>TNLP3</i>	TELEMAR	Telefonia fixa
<i>TNLP4</i>	TELEMAR	Telefonia fixa
<i>TOTS3</i>	Totvs	Tecnologia, internet e call-centers
<i>TOYB3</i>	TEC TOY	Consumo - brinquedos
<i>TOYB4</i>	TEC TOY	Consumo - brinquedos
<i>TPIS3</i>	Triunfo	Logística e transporte
<i>TRPL4</i>	CTEEP	Energia elétrica
<i>UGPA4</i>	Ultrapar	Investimentos
<i>UNIP6</i>	Unipar	Petroquímica
<i>UOLL4</i>	UOL	Tecnologia, internet e call-centers
<i>USIM3</i>	Usiminas	Siderurgia
<i>USIM5</i>	Usiminas	Siderurgia
<i>VALE3</i>	Vale	Mineração
<i>VALE5</i>	Vale	Mineração
<i>VIVO4</i>	VIVO	Telefonia móvel
<i>VLID3</i>	Valid	Serviços diversos
<i>WEGE3</i>	WEG	Máquinas e equipamentos

Table 3: stocks, companies, and sectors to which they belong.

B Sector distribution

In the following figures, I present the distribution of stocks according to sectors in three dimensional maps. One may see that stocks from the same sectors tend to agglomerate.

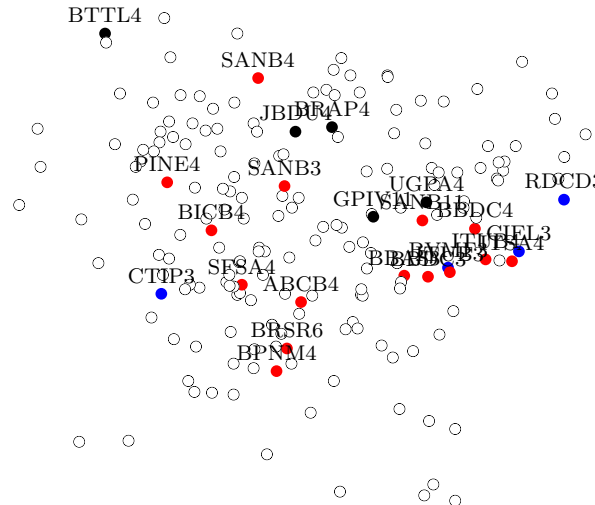


Figure 33: banking (red), financial (blue), and holdings (black) sectors.

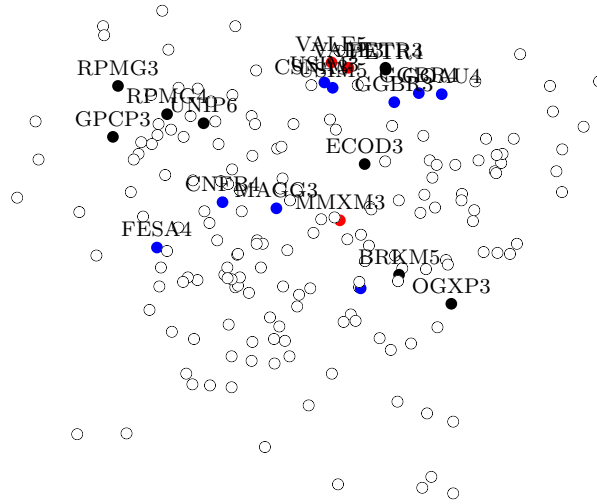


Figure 34: mining (red), ironworks (blue), and petroleum (black) sectors.

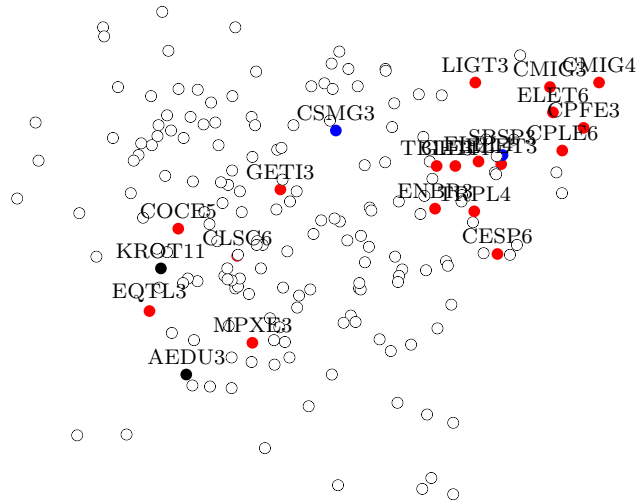


Figure 35: electricity (red), sanitation (blue), and education (black) sectors.

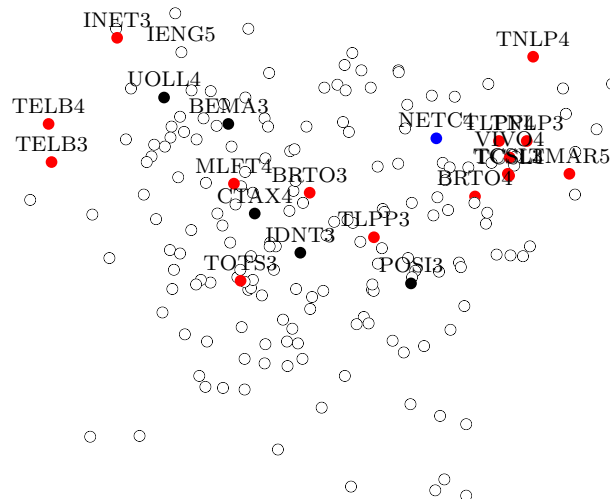


Figure 36: telecommunications (red), cable TV (blue), and technology, internet, and call-centers (black) sectors.

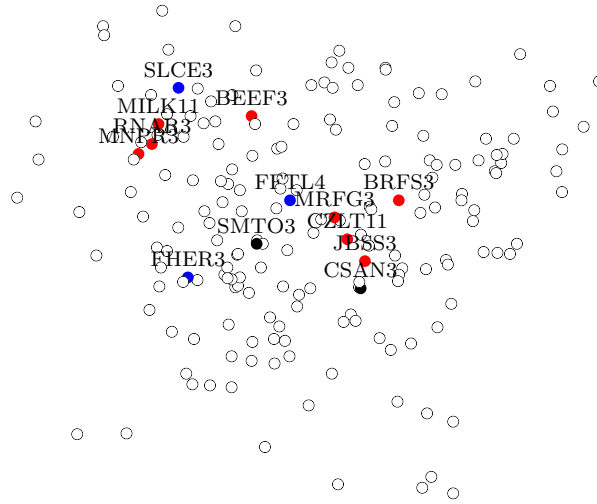


Figure 37: food (red), sugar and ethanol (blue), and agribusiness (black) sectors.

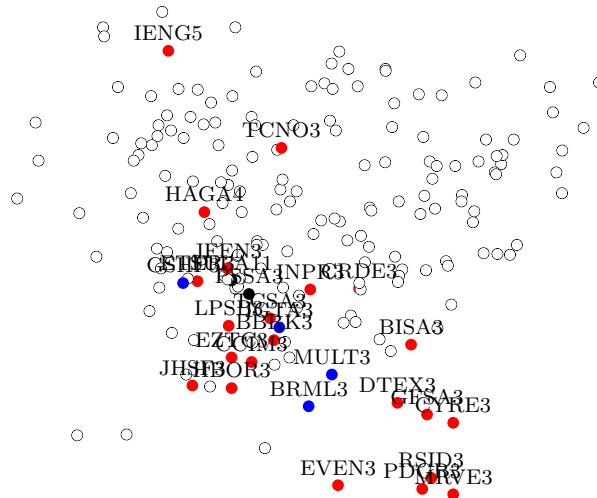


Figure 38: construction, materials, and engineering (red), real state (blue), and insurance (black).

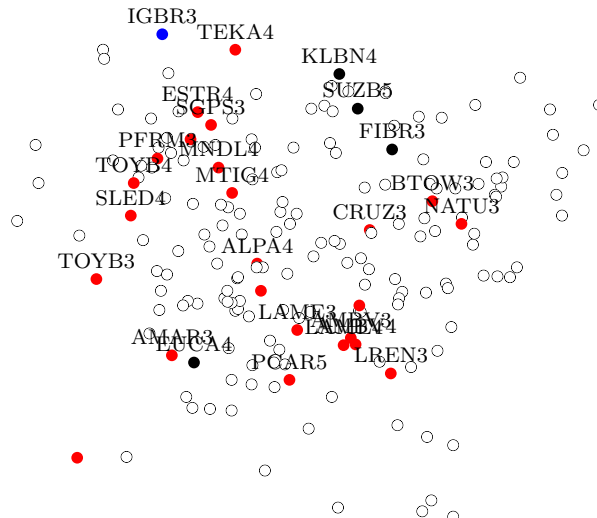


Figure 39: consumer goods (red), electronics (blue), and wood and paper (black).

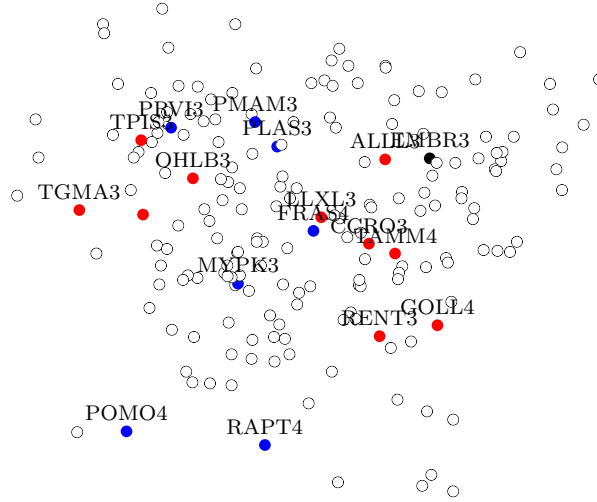


Figure 40: logistics and transportation (red), materials for transport and general (blue), airplane industry (black).

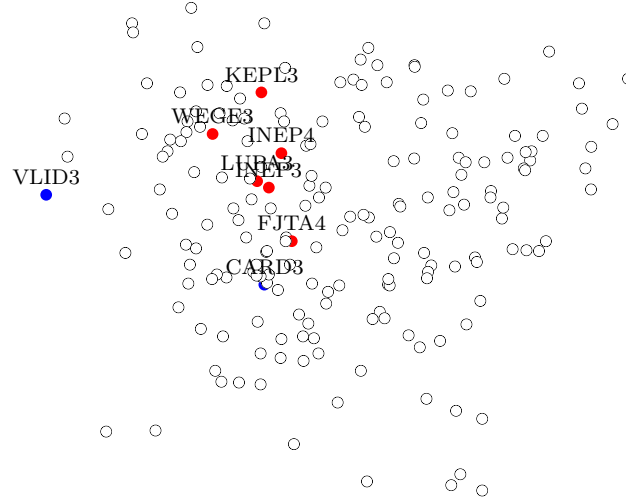


Figure 41: equipments and machinery (red), general services for payment (blue).

References

- [1] M. L. Mehta, *Random Matrices*, (2004) Academic Press.
- [2] L. Sandoval Jr. and I. De P. Franca, *Correlation of financial markets in times of crisis*, (2011) arXiv:1102.1339.
- [3] R.N. Mantegna, *Hierarchical structure in financial markets*, The European Phys. J. B **11** (1999) 193.
- [4] G. Bonanno, F. Lillo, and R.N. Mantegna, *High-frequency cross-correlation in a set of stocks*, Quantitative Finance **1** (2001) 96-104.
- [5] S. Micchichè, G. Bonanno, F. Lillo, and R.N. Mantegna, *Degree stability of a minimum spanning tree of price return and volatility*, Physica A **324** (2003) 66-73.
- [6] J.-P. Onnela, A. Chakraborti, K. Kaski, and J. Kertész, *Dynamic asset trees and Black Monday*, Physica A **324** (2003) 247-252.

- [7] A. Chakraborti, *An outlook on correlations in stock prices*, Econophysics of Stock and other Markets - Proceedings of the Econophys-Kolkata II (2007) Springer Milan.
- [8] J.-P. Onnela, A. Chakraborti, and K. Kaski, *Dynamics of market correlations: taxonomy and portfolio analysis*, Phys. Rev. E **68** (2003) 1-12.
- [9] J.-P. Onnela, A. Chakraborti, K. Kaski, J. Kertész, and A. Kanto, *Asset trees and asset graphs in financial markets*, Phys. Scripta T **106** (2003) 48-54.
- [10] J.-P. Onnela, K. Kaski, and J. Kertész, *Clustering and information in correlation based financial networks*, Eur. Phys. J. B **38** (2004) 353-362.
- [11] G. Bonanno, G. Caldarelli, F. Lillo, S. Micciché, N. Vandewalle, and R.N. Mantegna, *Networks of equities in financial markets*, The European Phys. J. B **38**, (2004) 363-371.
- [12] C. Coronello, M. Tumminello, F. Lillo, S. Micciché, and R.N. Mantegna, *Economic sector identification in a set of stocks traded at the New York Stock Exchange: a comparative analysis*, Proceedings of the SPIE, vol. 6601, 66010T (2007).
- [13] M. Tumminello, T. Di Matteo, T. Aste, and R.N. Mantegna, *Correlation based networks of equity returns sampled at different time horizons*, Eur. Phys. J. B **55** (2007) 209-217.
- [14] C. Borghesi, M. Marsili, and S. Micciché, *Emergence of time-horizon invariant correlation structure in financial returns by subtraction of the market mode*, Phys. Rev. E **76** (2007), 026104.
- [15] J.G. Brida and W.A. Risso, *Multidimensional minimal spanning tree: the Dow Jones case*, Physica A **387** (2008) 5205-5210.
- [16] Y. Zhanga, G.H.T. Leea, J.C. Wonga, J.L. Kokb, M. Prustyb, and S.A. Cheong, *Will the US Economy Recover in 2010? A Minimal Spanning Tree Study*, (2010) Physica A **390** (2011) 2020-2050.
- [17] M. Tumminello, F. Lillo, and R.N. Mantegna, *Correlation, hierarchies, and networks in financial markets*, Journal of Economic Behavior & Organizations (2010) Article in press.
- [18] S. Drożdż, J. Kwapień, and J. Speth, *Coherent patterns in nuclei and in financial markets*, AIP Conf. Proc. **1261** (2010), 256-264.
- [19] Y.W. Goo, T.W. Lian, W.G. Ong, W.T. Choi, and S.A. Cheong, *Financial atoms and molecules*, preprint (2009) arXiv:0903.2009v1.
- [20] C. Coronello, M. Tumminello, F. Lillo, S. Micciché, and R.N. Mantegna, *Sector identification in a set of stock return time series traded at the London Stock Exchange*, Acta Phys. Pol. B **36** (2005) 2653-2679.
- [21] S.A. Cheong, R.P. Forna, G.H.T. Lee, J.L. Kok, W.S. Yim, D.Y. Xu, Y. Zhang, *The Japanese Economy in Crises: A Time Series Segmentation Study*, Economics **2011-24** (2011)
- [22] S. Sinha and R.K. Pan, *Uncovering the internal structure of the Indian financial market: cross-correlation behavior in the NSE*, in "Econophysics of markets and business networks", Springer (2007) 215-226.
- [23] R. Coelho, C.G. Gilmore, B. Lucey, P. Richmond, and S. Hutzler, *The evolution of interdependence in world equity markets - evidence from minimum spanning trees*, Physica A **376** (2007) 455-466.
- [24] M. Ausloos and R. Lambiotte, *Clusters or networks of economies? A macroeconomy study through gross domestic product*, Physica A **382** (2007) 16-21.
- [25] C. Eom, G. Oh, W.-S. Jung, H. Jeong, and S. Kim, *Topological properties of stock networks based on minimal spanning tree and random matrix theory in financial time series*, Physica A **388** (2009), 900-906.

- [26] M. Eryiğit and R. Eryiğit, *Network structure of cross-correlations among the world market indices*, Physica A **388** (2009), 3551-3562.
- [27] P. Sieczka and J.A. Hołyst, *Correlations in commodity markets*, Physica A **388** (2009), 1621-1630.
- [28] J. Kwapien, S. Gworek, S. Drożdż, and A. Górski, *Analysis of a network structure of the foreign currency exchange market*, J. of Economic Interaction and Coordination **4** (2009), 55-72.
- [29] J. Kwapien, S. Gworek, and S. Drożdż, *Structure and evolution of the foreign exchange networks*, Acta Physica Polonica B **40** (2009), 175-194.
- [30] D.J. Fenn, M.A. Porter, P.J. Mucha, M. McDonald, S. Williams, N.F. Johnson, and N.S. Jones, *Dynamical clustering of exchange rates*, preprint (2010) arXiv:0905.4912v2.
- [31] M. Keskin, B. Deviren, and Y. Kocalkaplan, *Topology of the correlation networks among major currencies using hierarchical structure methods*, Physica A **390** (2011), 719-730.
- [32] G. Fagiolo, *The international-trade network: gravity equations and topological properties*, J. of Economic Interaction and Coordination **5** (2010), 1-25.
- [33] J. He and M.W. Deem, *Structure response in the world trade network*, Phys. Rev. Lett. **105** (2010), 198701-1 to 198701-4.
- [34] L. De Benedictis and L. Tajoli, *The World Trade Network* (2009), forthcoming in The World Economy.
- [35] K.M. Lee, J.-S. Yang, J. Lee, K.-I. Goh, and I.- M. Kim, *Impact of the topology of global macroeconomic network on the spreading of economic crises*, (2011)
PLoS ONE 6(3): e18443.doi:10.1371/journal.pone.0018443
- [36] E.J. Elton, M.J. Gruber, S.J. Brown, and W. Goetzmann, *Modern Portfolio Theory and Investment Analysis*, Eighth Edition , (2009) Wiley.
- [37] M.E.J. Newman, *Networks, and introduction*, (2010) Oxford University Press.